Përdorimi i mjeteve dhe metodave të Big Data për të shqyrtuar tendencat e çmimeve të banesave në Shqipëri

ZBATIM ME PYTHON DHE FUNKSIONALITETET ‘POWER’ TË MS EXCEL

Përgatiti: Orion Garo,
Departamenti i Kërkimeve, Banka e Shqipërisë.

Hyrje

Pronat rezidenciale janë një ‘niche’ i tregut të pasurive të paluajtshme në Shqipëri. Ky i fundit është një nga tregjet themelore të ekonomisë vendase. Pronat rezidenciale janë një segment i tij i përqendruar vetëm tek banesat, duke përjashtuar të tjera prona të paluajtshme si tokë e punueshme, pyje dhe kullota, hapësira tregtie ose biznesi, ndërtesa industriale, etj.

Aktualisht nuk ka statistika zyrtare mbi tregun e banesave në Shqipëri. Megjithatë çmimet e pasurive të paluajtshme, së bashku me datën e listimit, foto, dhe informacion tjetër përkatës, mund të gjenden të publikuara rregullisht në internet për informim të publikut të gjerë.

Një website i tillë i listimit të çmimeve është njoftime.com. Aty publikohen listime të reja çdo ditë, dhe arkivohen listimet për një periudhë deri në 5 vjet. Listimet e publikuara në këtë website (rrth 40.000 të tilla që datojnë nga janar 2017 to shkurt 2022) janë përdorur për këtë projekt kërkimor.

Hyrje – fokusi i projektit kërkimor

Përqendrimi tek ‘Big Data’:

- “Big Data” tipikisht i referohet volumit të pamatë të informacionit të hedhur në internet, sidomos UGC (përmbajtjeve të gjeneruara nga përdoruesi).
- Malli më me vlerë i kohëve aktuale është INFORMACIONI. Ai gjendet ngado; sidomos që nga shpërthimi i World Wide Web, ne jemi të rrethuar nga, çfarë quhet, ‘Big Data’;
- Sfida sot është të jesh në gjendje:
  a) të identifikosh informacion me vlerë në tërësine e ‘Big Data’ të panevojshme,
  b) ta mbledhësh, dhe
  c) ta përpunosh në një mënyrë sistematike (që më pas të të shërbejë për përdorim tradicional).

Motivimi për këtë projekt kërkimor:

- Tendencat për ecuritë e tregut shqiptar të banesave janë të vështira për t’u gjurmuar apo vlerësuar. Aktualisht nuk ka statistika periodike zyrtare, lidhur me çmimet për metër katror të ndonjë (cilësdo) pasurie të paluajtshme, që të publikohet nga Instituti i Statistikave apo ndonjë institucion tjetër publik.

Hyrje – objektivi

Objektivi:

- Vlerësimi i një indeksi mujor të çmimeve për metër katror për banesat ne shitje në kryeqytetin shqiptar.
  - *Ky objektiv synohet të arrihet duke:*
    a) *përdorur çmimet e listuara në internet për banesat në shitje në Tiranë, si dhe në zonat bregdetare, dhe territorin qendror të Shqipërisë;*
    b) *Vënë në zbatim mjetet dhe metodat e Big Data për sigurimin dhe organizimin e të dhënave – konkretisht: Python dhe funksionalitetet 'Power' të MS Excel – në mënyrë për të mundësuar identifikimin, mbledhjen, pastrimin, organizimin dhe llogaritjen e këtyre të dhënave të gjendura në website, me qëllim sintetizimin e indeksit të çmimit për të cilin interesohemi.*

Rishikimi i literaturës

Çmimet e banesave tipikisht karakterizohen nga variacione të mëdha, të cilat janë të pranishme për shkak të një morie faktorësh, të cilët përfshijnë:

- llojin dhe materialin e ndërtimit, amortizimin e ndërtesës, vendndodhjen e pronës,
- pozicionimin e katit, klimën, orientimin e ndërtesës,
- lagjen/fqinjësinë ku ndodhet prona, afërsinë me vende specifike, sipërfaqen e pronës,
- arranzhimin arkitektural të pronës, qëllimin e përdorimit të pronës, gatishmërinë e shitësit,

…etj.

Krahas tyre, shitblerjet e banesave në Shqipëri i nënshtrohen një taksimi specifik që pronarët (shitësit e pronës) duhet ta mbulojnë. Kjo taksë e shitjes së pronës llogaritet duke konsideruar një çmim reference që i atribuohet vendndodhjes gjeografike të pronës në shitje, dhe përcaktohet nga Ministria e Financave. Taksa mbi fitimin e siguruar nga shitja, madje dhe vetë çmimi i referencës, mund të influencojnë ujdinë midis shitësit dhe blerësit mbi çmimin përfundimtar të pronës në shitje.

Për më tepër, një pjesë e konsiderueshme e shitësve tipikisht blejnë shërbimet e një agjencie pronash të paluajtshme në mënyrë që t'ja reklamojnë shitjen një mase sa më të gjerë blerësish potencialë, dhe këto shërbime kanë tendencë ta rrisin /(influencojnë mbi) çmimin përfundimtar të shitjes.

Rishikimi i literaturës

Të dhëna statistikore mbi tregun e banesave të shkarkuara nga website-t po përdoren gjithnjë e më shpesh në lëmin e kërkimit ekonomik; vëmë re këtu punimet kërkimore nga:

1. Chapelle & Eymeoud (2018), dhe
2. Meunier & Pouget (2021) të *Bankës së Francës*,

...me qëllim për të siguruar njohuri më të thelluara mbi tendencat e tregut të banesave, qoftë në tregun e qirave apo në atë të banesave në shitje.

Ky lloj të dhënash:

a) ofron perspektivën e shitësit, nëpërmjet listimeve që këta të fundit, ose agjencitë e pronave të paluajtshme që këta kontraktojnë, postojnë në internet (Bricongne et al, 2021);

b) publikohet menjëherë; çmimet e listuara të pronave jepen në kohë reale, duke bërë që tendencat e tregut të kapen praktikisht pa vonesa;

c) Ofron epërsi – krahasuar me të dhënat nga sondazhet – në lidhje me kostot financiare, numrin e vrojtimeve, mbulimin gjegrafik dhe ‘granularitetin’ (Chapelle & Eymeoud, 2018).

Të dhënat dhe kuadri metodologjik

Metodat e mbledhjes së të dhënave, të zbatuara me programin Python, që përdoren në këtë projekt kërkimor përshkruhen në Hajba G. L. (2018)[1] dhe Kouzis-Loukas D. (2016)[2].

Në thelb, procedura kërkon: [i] brenda mjedisit të Python, instalimin, konfigurimin dhe përdorimin e motorit të shkarkimit të të dhënave të quajtur Scrapy, i ndërtuar për të kryer shkarkimin elektronik të të dhënave; [ii] identifikimin e website-t që ka të dhënat, dhe; [iii] futjen në strukturën e tij HTML për të lokalizuar copëzat e nevojshme të informacionit të cilat planifikohet të ruhen për çdo çmim të listuar (vrojtim).

Në rastin tonë, 4 copëzat e informacionit që duhet të lokalizohen dhe ruhen, për çdo çmim të listuar, janë:

I. Data (e listimit për shitje të pronës);
II. Agjenti listues (ai/ajo që kryen shtimin e listimit për shitje);
III. Kreu (detajet më të rëndësishme mbi pronën, si çmimi i listuar dhe sipërfaqja në metër katrorë);
IV. Informacion (detaje më të zgjeruara mbi pronën, si lokalizimi, kati, aranzhimi arkitektural, numri i dhomave, etj.).

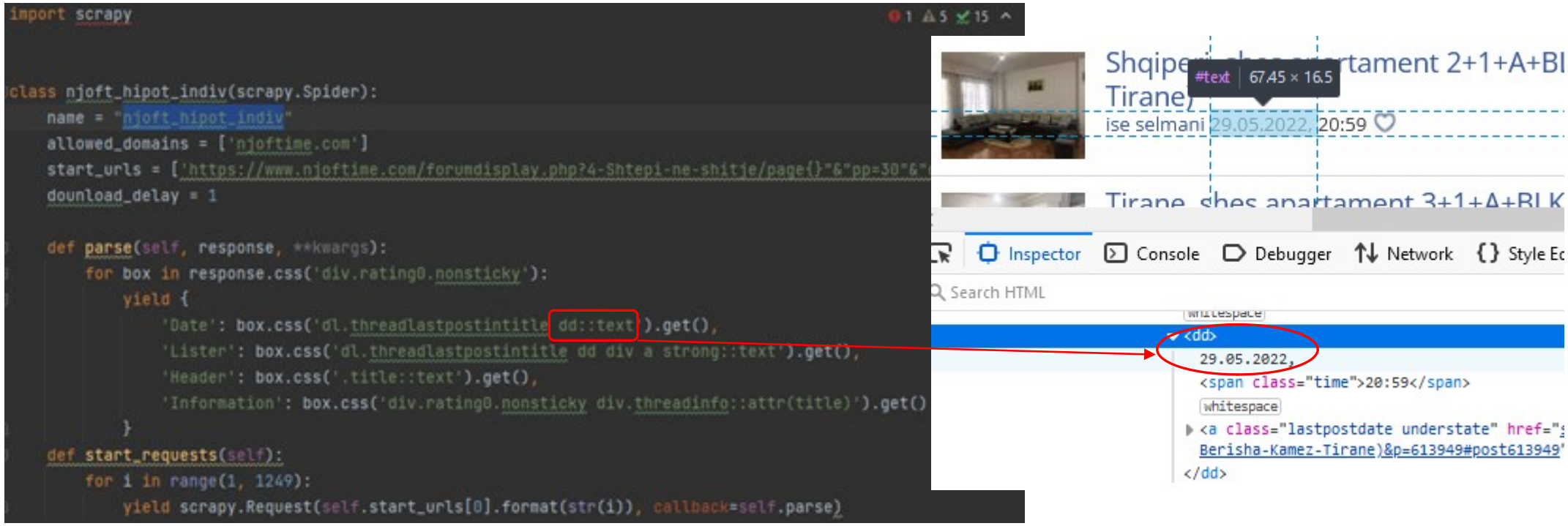
! Tirane,shes Apartament me hipoteke 2+1+A+BLK Kati 2, 107 m²,160.000 Euro(SECRET GARDEN-TEG)
EuroEstate Albania 11.04.2022, 20:15
Zona/Rruga : TEG
Apartamenti ndodhet ne nje nga rezidencat me te bukura ne Tirane, prane TEG-ut.
Sip totale eshte 107 m2,me ashensor.
Punime super cilesore ,lartesi 3 mt ,veshje me gur,izolim akustik etj
Eshte i organizuar ne :Sallon, 2...

1. *Hajba, Gábor László - Website scrapping with Python, 2018.*
2. *Kouzis-Loukas, Dimitrios - Learning Scrapy: Learn the art of efficient web scraping and crawling with Python, 2016.*

Të dhënat dhe kuadri metodologjik

Shembull:

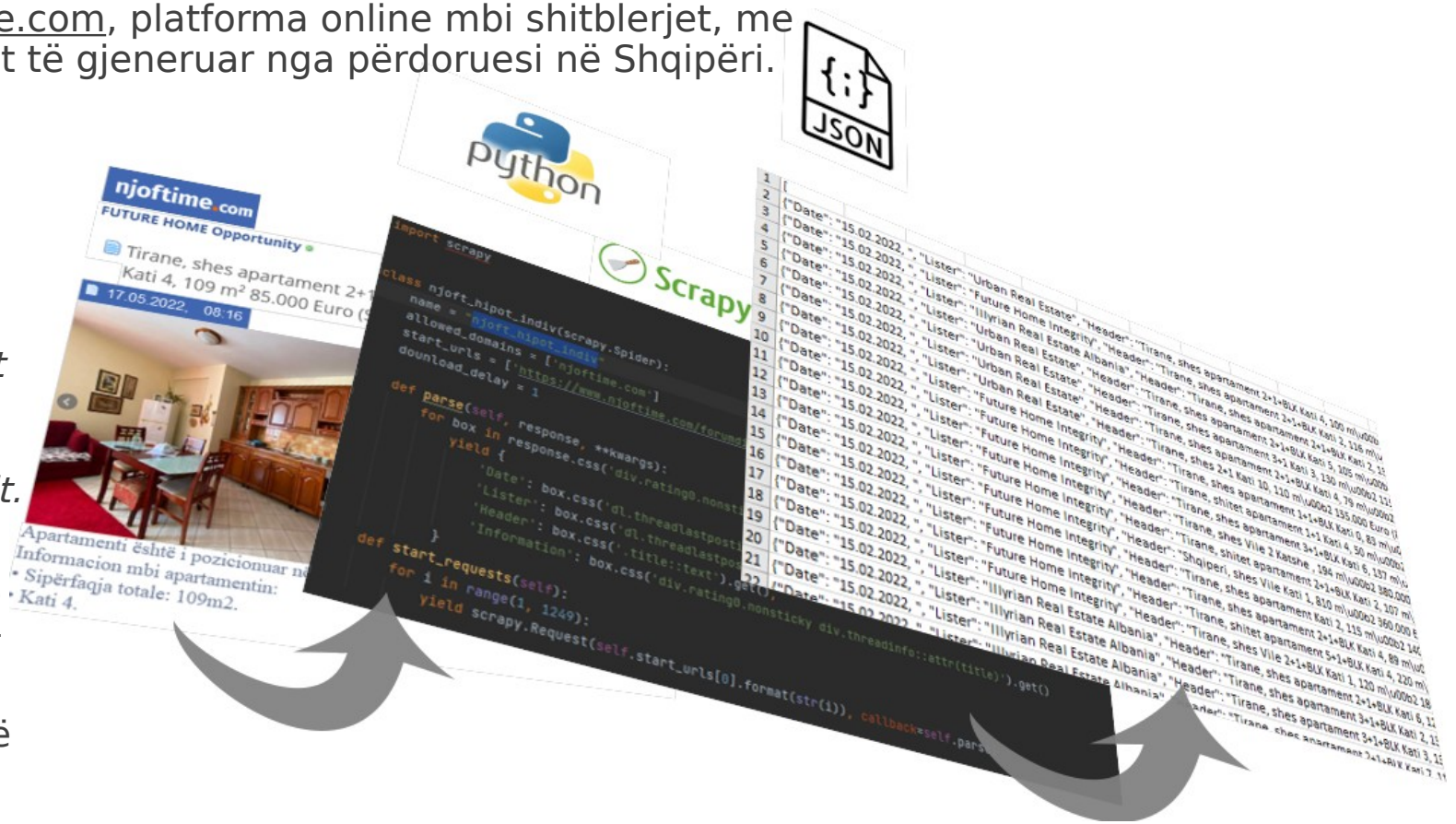
Në të majtë: Një kod komande në Python, për udhëzimin e Scrapy për të kryer shkarkimin e të dhënave nga website i synuar; *në të djathtë*: struktura HTML e website (nga ku shkarkohen të dhënat).



Të dhënat dhe kuadri metodologjik

Të dhënat janë marrë nga njoftime.com, platforma online mbi shitblerjet, me bazën më të gjerë të informacionit të gjeneruar nga përdoruesi në Shqipëri.

- U shkarkuan 40.491 listime çmimesh të banesave në shitje, nga janar 2017 deri në shkurt 2022. Nga keto:
  - *87.7% ishin çmime për shitje vetëm në Tiranë;*
  - *8.3% çmime për shitje në zonat bregdetare;*
  - *3.9% çmime për shitje në territorin e brendshëm të vendit.*
- Python, dhe nën-komponenti i tij Srapy, u përdorën për të marrë të dhënat nga përmbajtja e website-it.
- Procedura e përdorur gjeneron një file *.json*, ku ruhen të dhënat e papërpunuara.

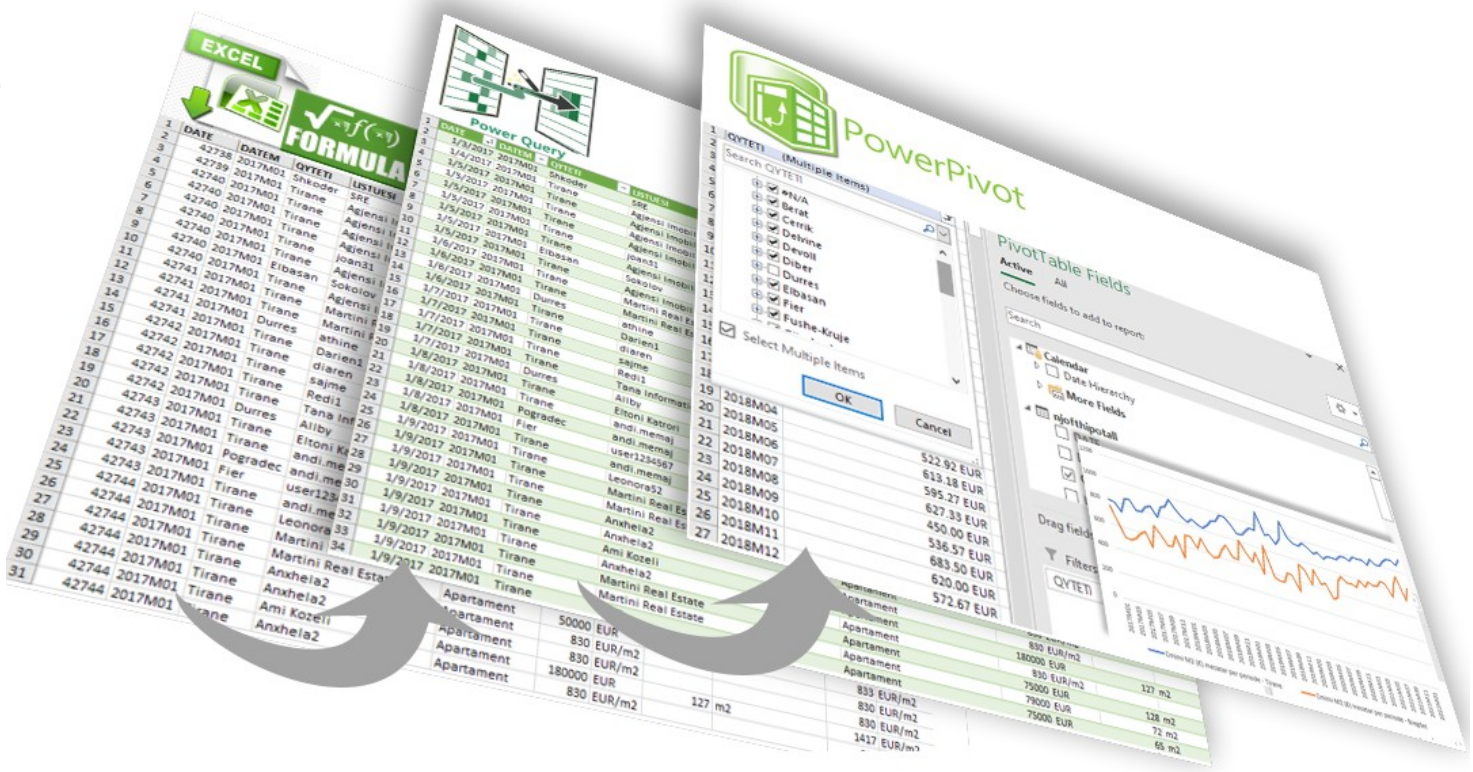


Të dhënat dhe kuadri metodologjik

Me t'u ruajtur të dhënat e papërpunuara në filen *.json*, me anë të formulave Excel kryhet filtrimi, pastrimi dhe organizimi i tyre në formë tabulare, me variabla të specifikuara, pa

Kur setet e të dhënave në formë tabulare janë gati, nëpërmjet PowerQuery, ata bëhen bashkë (i mbivendosen njeri-tjetrit) në një kuadër/model të vetëm të dhënash.

Ky model të dhënash më pas i nënshtrohet përpunimit në PowerPivot dhe gjuhën DAX për të sintetizuar rezultatet e nevojitura (siguruar setet, raportet dhe grafikët e filtruar, për frekuencën kohore të specifikuar).



Të dhënat dhe kuadri metodologjik

Specifika për t'u sqaruar në lidhje me produktin përfundimtar (të dhënat tabulare të organizuara):

a) Frekuenca kohore është caktuar mujore – mund të ishte caktuar tremujore, javore, madje dhe ditore (në varësi të mjaftueshmërisë së vrojtimeve);

b) Rezultojnë rreth 648 vrojtime për periudhë (muaj);

c) Të dhënat mbi çmimet e shitjeve të banesave në këtë projekt kërkimor janë ndarë – duke ju referuar metodologjisë së përdorur në pyetësorin mbi "Ecurinë e Tregut të Pasurive të Paluajtshme në Shqipëri", të kryer nga Departamenti i Stabilitetit Financiar, i Bankës së Shqipërisë – në 3 grupe:

1. Tirana (përfshihet bashkia Tiranë)

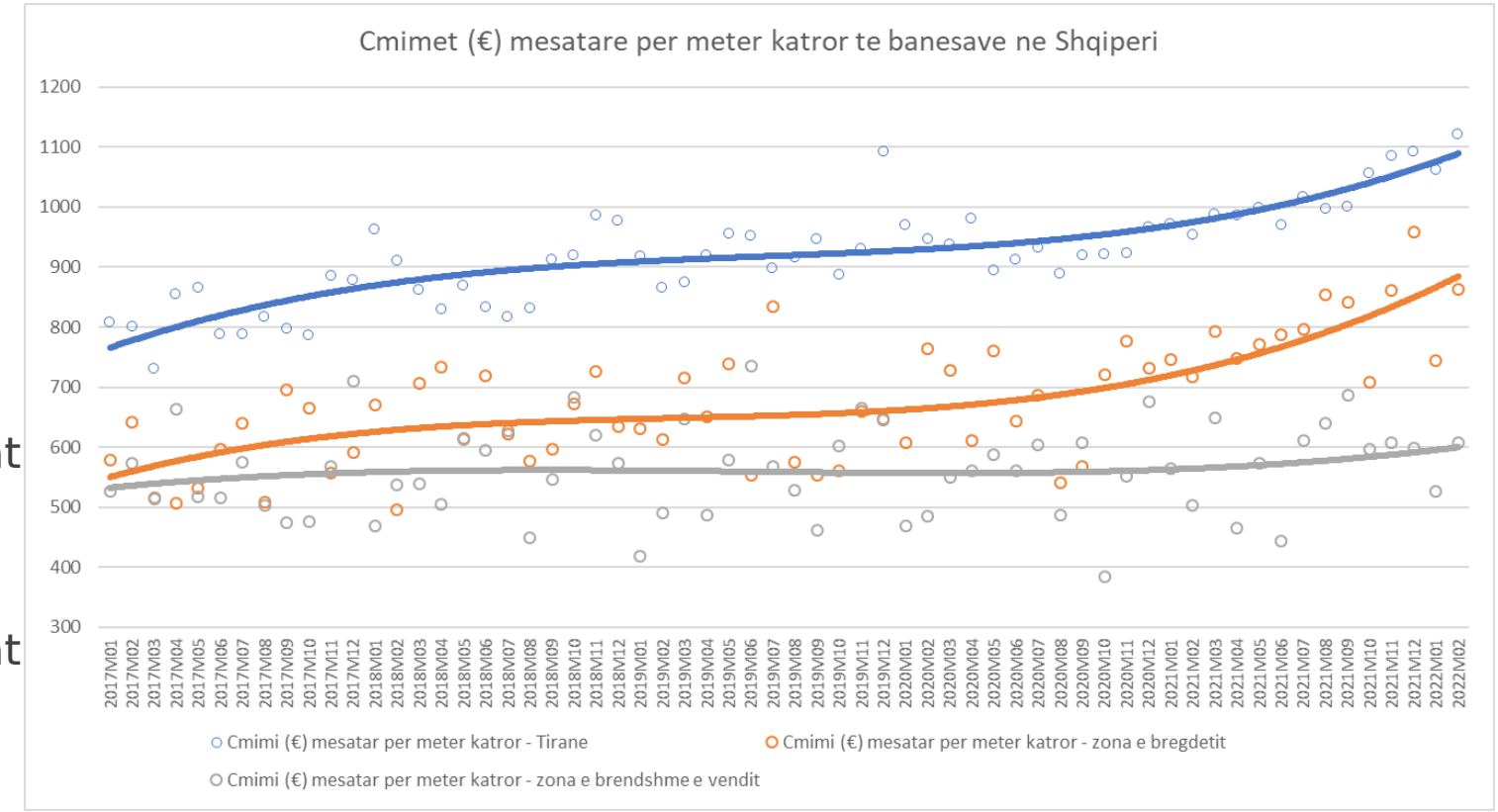
2. Qytetet, qytezat dhe zonat rurale bregdetare të Shqiprërisë (përfshihen bashkitë):
   Durrës, Himarë, Kavajë, Sarandë, Shëngjin, Vlorë.

3. Qytetet, qytezat dhe zonat rurale të territorit të brendshëm të Shqipërisë (përfshihen bashkitë):
   Berat, Cërrik, Delvinë, Devoll, Dibër, Elbasan, Fier, Fushë-Krujë, Gjirokastër, Gramsh, Kamëz, Korçë, Krujë, Kuçovë, Kukës, Laç, Lezhë, Librazhd, Lushnjë, M. e Madhe, Mirditë, Pogradec, Përmet, Shkodër, Tepelenë.

Rezultatet paraprake

Grafiku tregon mesataret e cmimeve për metër katror në Euro, për 3 kategoritë specifike të lokaliteteve të banesave[3]:

1. ato të lokalizuara brenda kufinjve të Bashkisë Tiranë (blu);
2. ato të lokalizuara në qytetet, qytezat dhe zonat rurale bregdetare të Shqiprërisë (portokalli);
3. ato të lokalizuara në qytetet, qytezat dhe zonat rurale të territorit të brendshëm të Shqipërisë (gri).



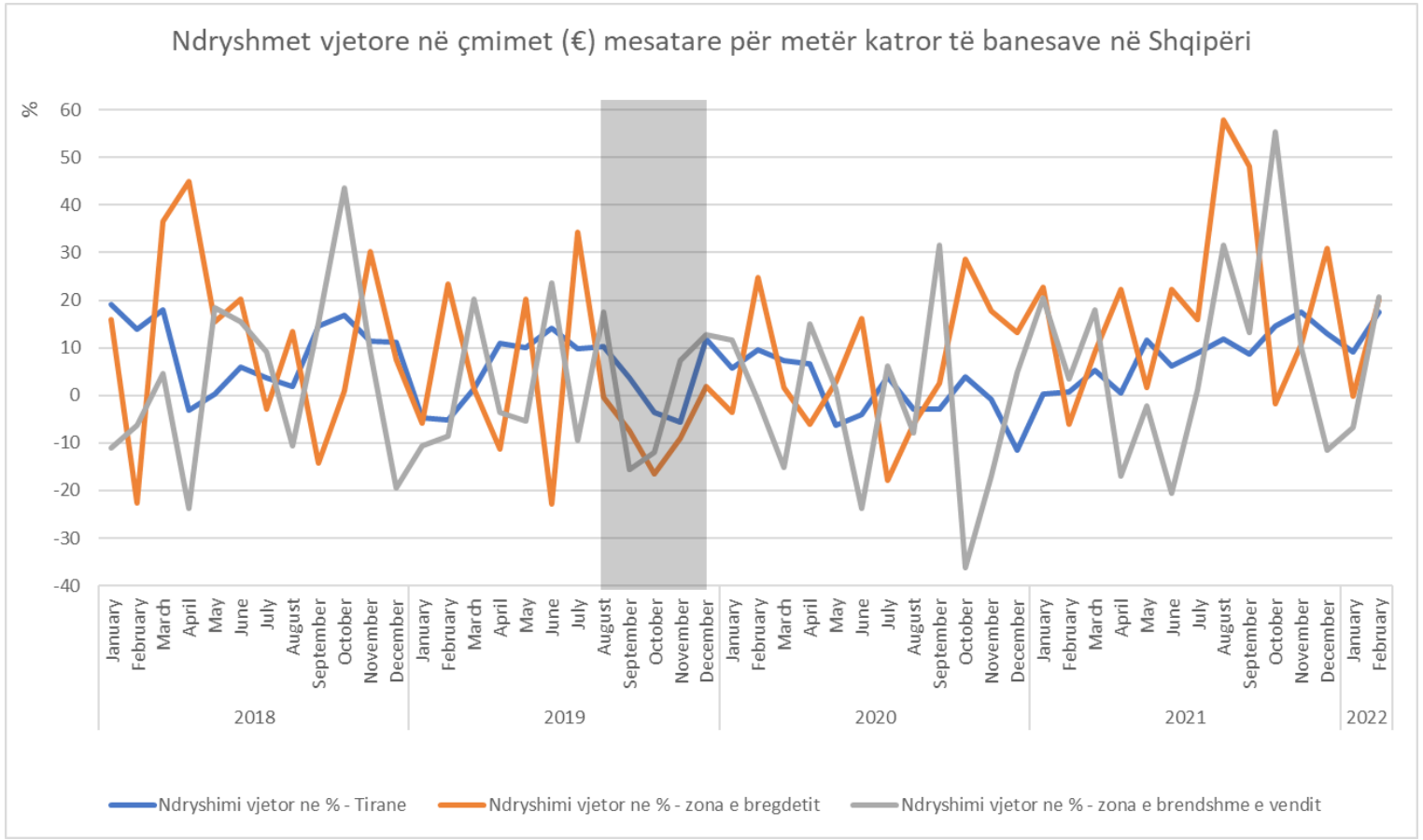
Rezultatet paraprake

Mesataret e cmimeve, dhe ndryshmet vjetore (YoY) korresponduese në përqindje: për vitin 2021, dhe muajin e fundit për të cilin ka të dhëna (shkurt 2022), jepen në tabelën e mëposhtme:

	Mesatarja e cmimeve për 2021 (€)	Ndryshmi vjetor (YoY) në përqindje i mesatares së cmimeve për 2021	Mesatarja e cmimeve për shkurt 2022 (€)	Ndryshmi vjetor (YoY) në përqindje i mesatares së cmimeve për shkurt 2022
Bashkia Tiranë	1,009.86	8.31%	1,121.65	17.54%
Qytetet, qytezat dhe zonat rurale bregdetare të Shqiprërisë	793.52	19.01%	862.8	20.25%

Rezultatet paraprake

- Ndryshmet vjetore (YoY) në përqindje të mesatareve të cmimeve shfaqin luhatje të theksuara për zonën bregdetare dhe territorin qendror të Shqipërisë.
  - Kjo për shkak se përqindja e nr. të vrojtimeve për këto zona është e vogël krahasuar me Tiranën (kjo e fundit e ka 87.7%)
  - (Për t'u kryer, kontrollet për sezonalitet).
- Të 3 treguesit shfaqin rritje negative gjatë gjysmës së dytë të 2019 *(fatkeqësia natyrore e tërmeteve);*
- Nga fillimi i 2021 e në vijim, rritjet e treguesve vërehen të marrin të treja trend pozitiv.



Faleminderit për vëmendjen!

Vlerësoj pyetjet, observimet, komentet!

Përdorimi i mjeteve dhe metodave të Big Data për të shqyrtuar tendencat e çmimeve të banesave në Shqipëri

ZBATIM ME PYTHON DHE FUNKSIONALITETET 'POWER' TË MS EXCEL